

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT: Fuerst, Thomas R.
 McAtee, C. Patrick
 Yarbough, Patrice O.
 Zhang, Yifan

10 (ii) TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

 (iii) NUMBER OF SEQUENCES: 31

 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Dehlinger & Associates
 (B) STREET: 350 Cambridge Ave., Suite 250
 (C) CITY: Palo Alto
 (D) STATE: CA
 (E) COUNTRY: USA
20 (F) ZIP: 94306

 (v) COMPUTER READABLE FORM:

 (A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

 (vi) CURRENT APPLICATION DATA:

 (A) APPLICATION NUMBER:
30 (B) FILING DATE:
 (C) CLASSIFICATION:

 (viii) ATTORNEY/AGENT INFORMATION:

 (A) NAME: Fabian, Gary R.
35 (B) REGISTRATION NUMBER: 33,875
 (C) REFERENCE/DOCKET NUMBER: 4600-0293.30

 (ix) TELECOMMUNICATION INFORMATION:

 (A) TELEPHONE: (415) 324-0880
40 (B) TELEFAX: (415) 324-0960

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2049 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
 ORF-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

20	ATGCGCCCTC GGCCTATTTT GTTGCTGCTC CTCATGTTTT TGCCTATGCT GCCCGCGCCA	60
	CCGCCCGGTC AGCCGTCTGG CCGCCGTCGT GGGCGGCGCA GCGGCGGTTC CGGCGGTGGT	120
	TTCTGGGGTG ACCGGGTGA TTCTCAGCCC TTCGCAATCC CCTATATTCA TCCAACCAAC	180
25	CCCTTCGCCC CCGATGTCAC CGCTGCGGCC GGGGCTGGAC CTCGTGTTCG CCAACCCGCC	240
	CGACCACTCG GCTCCGCTTG GCGTGACCAG GCCCAGCGCC CCGCCGTTGC CTCACGTCGT	300
30	AGACCTACCA CAGCTGGGGC CGCGCCGCTA ACCGCGGTCG CTCCGGCCCA TGACACCCCG	360
	CCAGTGCCTG ATGTCGACTC CCGCGGCGCC ATCTTGCGCC GGCAGTATAA CCTATCAACA	420
	TCTCCCCTTA CCTCTTCCGT GGCCACCGGC ACTAACCTGG TTCTTTATGC CGCCCCCTCTT	480
35	AGTCCGCTTT TACCCCTTCA GGACGGCACC AATACCCATA TAATGGCCAC GGAAGCTTCT	540
	AATTATGCCC AGTACCGGGT TGCCCGTGCC ACAATCCGTT ACCGCCCGCT GGTCCCCAAT	600
40	GCTGTCGGCG GTTACGCCAT CTCCATCTCA TTCTGGCCAC AGACCACCAC CACCCCGACG	660

	TCCGTTGATA TGAATTCAAT AACCTCGACG GATGTTGTA TTTTAGTCCA GCCCGGCATA	720
	GCCTCTGAGC TTGTGATCCC AAGTGAGCGC CTACACTATC GTAACCAAGG CTGGCGCTCC	780
5	GTCGAGACCT CTGGGGTGGC TGAGGAGGAG GCTACCTCTG GTCTTGTTAT GCTTTGCATA	840
	CATGGCTCAC TCGTAAATTC CTATACTAAT ACACCCTATA CCGGTGCCCT CGGGCTGTTG	900
10	GACTTTGCCC TTGAGCTTGA GTTTCGCAAC CTTACCCCCG GTAACACCAA TACGCGGGTC	960
	TCCCGTTATT CCAGCACTGC TCGCCACCGC CTTGTCGCG GTGCGGACGG GACTGCCGAG	1020
	CTCACCACCA CGGCTGCTAC CCGCTTTATG AAGGACCTCT ATTTTACTAG TACTAATGGT	1080
15	GTCGGTGAGA TCGGCCGCGG GATAGCCCTC ACCCTGTTCA ACCTTGCTGA CACTCTGCTT	1140
	GGCGGCCTGC CGACAGAATT GATTTCGTCG GCTGGTGGCC AGCTGTTCTA CTCCCGTCCC	1200
20	GTTGTCTCAG CCAATGGCGA GCCGACTGTT AAGTTGTATA CATCTGTAGA GAATGCTCAG	1260
	CAGGATAAGG GTATTGCAAT CCCGCATGAC ATTGACCTCG GAGAATCTCG TGTGGTTATT	1320
	CAGGATTATG ATAACCAACA TGAACAAGAT CGGCCGACGC CTTCTCCAGC CCCATCGCGC	1380
25	CCTTCTCTG TCCTTCGAGC TAATGATGTG CTTTGGCTCT CTCTCACC GC TGCCGAGTAT	1440
	GACCAGTCCA CTTATGGCTC TTCGACTGGC CCAGTTTATG TTTCTGACTC TGTGACCTTG	1500
30	GTTAATGTTG CGACCGGCGC GCAGGCCGTT GCCCGGTCGC TCGATTGGAC CAAGGTCACA	1560
	CTTGACGGTC GCCCCCTCTC CACCATCCAG CAGTACTCGA AGACCTTCTT TGTCTGCCG	1620
	CTCCGCGGTA AGCTCTCTTT CTGGGAGGCA GGCACAATA AAGCCGGGTA CCCTTATAAT	1680
35	TATAACACCA CTGCTAGCGA CCAACTGCTT GTCGAGAATG CCGCCGGGCA CCGGGTCGCT	1740
	ATTTCCACTT ACACCACTAG CCTGGGTGCT GGTCCCGTCT CCATTTCTGC GGTGCGCGTT	1800
40	TTAGCCCCC ACTCTGCGCT AGCATTGCTT GAGGATACCT TGGACTACCC TGCCGCGCC	1860
	CATACTTTTG ATGATTCTG CCCAGAGTGC CGCCCCCTTG GCCTTCAGGG CTGCGCTTTC	1920

CAGTCTACTG TCGCTGAGCT TCAGCGCCTT AAGATGAAGG TGGGTAAAAC TCGGGAGTTG 1980
 TAGTTTATTT GCTTG TGCCC CCCTTCTTTC TGTTGCTTAT TTCTCATTTT TCGGTTCCGC 2040
 5 GCTCCCTGA 2049

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 2058 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

20 (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
 ORF-2 region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

25 ATGCGCCCTA GGCCTCTTTT GCTGTTGTTT CTCTTGTTTC TGCCTATGTT GCCCGCGCCA 60
 CCGACCGGTC AGCCGTCTGG CCGCCGTCGT GGGCGGCGCA GCGGCGGTAC CGGCGGTGGT 120
 TTCTGGGGTG ACCGGGTGA TTCTCAGCCC TCGCAATCC CCTATATTCA TCCAACCAAC 180
 30 CCCTTTGCCC CAGACGTTGC CGCTGCGTCC GGGTCTGGAC CTCGCCTTCG CCAACCAGCC 240
 CGGCCACTTG GCTCCACTTG GCGAGATCAG GCCCAGCGCC CCTCCGCTGC CTCCCGTCGC 300
 35 CGACCTGCCA CAGCCGGGGC TCGGCGCTG ACGGCTGTGG CGCCTGCCCC TGACACCTCA 360
 CCCGTCCCGG ACGTTGATTC TCGCGGTGCA ATTCTACGCC GCCAGTATAA TTTGTCTACT 420
 TCACCCCTGA CATCCTCTGT GGCCTCTGGC ACTAATTTAG TCCTGTATGC AGCCCCCTT 480
 40 AATCCGCCTC TGCCGCTGCA GGACGGTACT AATACTCACA TTATGGCCAC AGAGGCCTCC 540

AATTATGCAC AGTACCGGGT TGCCCGCGCT ACTATCCGTT ACCGGCCCCT AGTGCCTAAT 600
 GCAGTTGGAG GCTATGCTAT ATCCATTTCT TTCTGGCCTC AAACAACCAC AACCCCTACA 660
 5 TCTGTTGACA TGAATTCCAT TACTTCCACT GATGTCAGGA TTCTTGTTCA ACCTGGCATA 720
 GCATCTGAAT TGGTCATCCC AAGCGAGCGC CTTCACTACC GCAATCAAGG TTGGCGCTCG 780
 GTTGAGACAT CTGGTGTTCG TGAGGAGGAA GCCACCTCCG GTCTTGTCAT GTTATGCATA 840
 10 CATGGCTCTC CAGTTAACTC CTATACCAAT ACCCCTTATA CCGGTGCCCT TGGCTTACTG 900
 GACTTTGCCT TAGAGCTTGA GTTTCGCAAT CTCACCACCT GTAACACCAA TACACGTGTG 960
 15 TCCCGTTACT CCAGCACGGC CCGTCACCGG CTCCGCCGAG GGGCCGACGG GACTGCGGAG 1020
 CTGACCACAA CTGCAGCCAC CAGGTTTCATG AAAGATCTCC ACTTTACCGG CCTTAATGGG 1080
 GTAGGTGAAG TCGGCCGCGG GATAGCTCTA ACATTACTTA ACCTTGCTGA CACGCTCCTC 1140
 20 GGCGGGCTCC CGACAGAATT AATTCGTCG GCTGGCGGGC AACTGTTTTA TTCCCGCCCG 1200
 GTTGTCTCAG CCAATGGCGA GCCAACCGTG AAGCTCTATA CATCAGTGGA GAATGCTCAG 1260
 25 CAGGATAAGG GTGTTGCTAT CCCCCACGAT ATCGATCTTG GTGATTGCGG TGTGGTCATT 1320
 CAGGATTATG ACAACCAGCA TGAGCAGGAT CGGCCCCACCC CGTCGCCTGC GCCATCTCGG 1380
 CCTTTTTCTG TTCTCCGAGC AAATGATGTA CTTTGGCTGT CCCTCACTGC AGCCGAGTAT 1440
 30 GACCAGTCCA CTTACGGGTC GTCAACTGGC CCGGTTTATA TCTCGGACAG CGTGACTTTG 1500
 GTGAATGTTG CGACTGGCGC GCAGGCCGTA GCCCGATCGC TTGACTGGTC CAAAGTCACC 1560
 35 CTCGACGGGC GGCCCCCTCC GACTGTTGAG CAATATTCCA AGACATTCTT TGTGCTCCCC 1620
 CTTCTGTTGCA AGCTCTCCTT TTGGGAGGCC GGCACAACAA AAGCAGGTTA TCCTTATAAT 1680
 TATAATACTA CTGCTAGTGA CCAGATTCTG ATTGAAAATG CTGCCGGCCA TCGGGTCGCC 1740
 40 ATTTCAACCT ATACCACCAG GCTTGGGGCC GGTCCGGTCG CCATTTCTGC GGCCGCGGTT 1800

5
10

15

20

- 25

30

35

- 40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:						
GCGGTCGCTC	CGGCCCATGA	CACCCCGCCA	GTGCCTGATG	TCGACTCCCG	CGGCGCCATC	60
TTGCGCCGGC	AGTATAACCT	ATCAACATCT	CCCCTTACCT	CTTCCGTGGC	CACCGGCACT	120
AACCTGGTTC	TTTATGCCGC	CCCTCTTAGT	CCGCTTTTAC	CCCTTCAGGA	CGGCACCAAT	180
ACCCATATAA	TGGCCACGGA	AGCTTCTAAT	TATGCCCAGT	ACCGGGTTGC	CCGTGCCACA	240
ATCCGTTACC	GCCCCTGGT	CCCCAATGCT	GTCGGCGGTT	ACGCCATCTC	CATCTCATTC	300
TGGCCACAGA	CCACCACCAC	CCCGACGTCC	GTTGATATGA	ATTCAATAAC	CTCGACGGAT	360

	GTTCGTATTT TAGTCCAGCC CGGCATAGCC TCTGAGCTTG TGATCCCAAG TGAGCGCCTA	420
	CACTATCGTA ACCAAGGCTG GCGCTCCGTC GAGACCTCTG GGGTGGCTGA GGAGGAGGCT	480
5	ACCTCTGGTC TTGTTATGCT TTGCATACAT GGCTCACTCG TAAATTCCTA TACTAATACA	540
	CCCTATACCG GTGCCCTCGG GCTGTTGGAC TTTGCCCTTG AGCTTGAGTT TCGCAACCTT	600
10	ACCCCCGGTA ACACCAATAC GCGGGTCTCC CGTTATTCCA GCACTGCTCG CCACCGCCTT	660
	CGTCGCGGTG CGGACGGGAC TGCCGAGCTC ACCACCACGG CTGCTACCCG CTTTATGAAG	720
	GACCTCTATT TTACTAGTAC TAATGGTGTC GGTGAGATCG GCCGCGGGAT AGCCCTCACC	780
15	CTGTTCAACC TTGCTGACAC TCTGCTTGGC GGCCTGCCGA CAGAATTGAT TTCGTCGGCT	840
	GGTGGCCAGC TGTTCTACTC CCGTCCCGTT GTCTCAGCCA ATGGCGAGCC GACTGTTAAG	900
20	TTGTATACAT CTGTAGAGAA TGCTCAGCAG GATAAGGGTA TTGCAATCCC GCATGACATT	960
	GACCTCGGAG AATCTCGTGT GGTTATTCAG GATTATGATA ACCAACATGA ACAAGATCGG	1020
	CCGACGCCTT CTCCAGCCCC ATCGCGCCCT TTCTCTGTCC TTCGAGCTAA TGATGTGCTT	1080
25	TGGCTCTCTC TCACCGCTGC CGAGTATGAC CAGTCCACTT ATGGCTCTTC GACTGGCCCA	1140
	GTTTATGTTT CTGACTCTGT GACCTTGGTT AATGTTGCGA CCGGCGCGCA GGCCGTTGCC	1200
30	CGGTCGCTCG ATTGGACCAA GGTCACACTT GACGGTCGCC CCCTCTCCAC CATCCAGCAG	1260
	TACTCGAAGA CCTTCTTTGT CCTGCCGCTC CGCGGTAAGC TCTCTTTCTG GGAGGCAGGC	1320
	ACAACTAAAG CCGGGTACCC TTATAATTAT AACACCACTG CTAGCGACCA ACTGCTTGTC	1380
35	GAGAATGCCG CCGGGCACC GGTGCTATT TCCACTTACA CCACTAGCCT GGGTGCTGGT	1440
	CCCGTCTCCA TTTCTGCGGT TGCCGTTTGA GCCCCCACT CTGCGCTAGC ATTGCTTGAG	1500
40	GATACCTTGG ACTACCCTGC CCGCGCCCAT ACTTTTGATG ATTTCTGCCC AGAGTGCCGC	1560
	CCCCTTGGCC TTCAGGGCTG CGCTTTCAG TCTACTGTCG CTGAGCTTCA GCGCCTTAAG	1620

ATGAAGGTGG GTAAACTCG GGAGTTG

1647

(2) INFORMATION FOR SEQ ID NO:4:

5 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1647 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
r62kDa, FIGURE 2

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCTGTGGCGC CTGCCCATGA CACCTCACCC GTCCCGGACG TTGATTCTCG CGGTGCAATT	60
CTACGCCGCC AGTATAATTT GTCTACTTCA CCCCTGACAT CCTCTGTGGC CTCTGGCACT	120
AATTTAGTCC TGTATGCAGC CCCCCTTAAT CCGCCTCTGC CGCTGCAGGA CGGTACTAAT	180
ACTCACATTA TGGCCACAGA GGCCTCCAAT TATGCACAGT ACCGGGTGTC CCGCGCTACT	240
ATCCGTTACC GGCCCCTAGT GCCTAATGCA GTTGGAGGCT ATGCTATATC CATTTCTTTC	300
TGGCCTCAAA CAACCACAAC CCCTACATCT GTTGACATGA ATTCCATTAC TTCCACTGAT	360
GTCAGGATTG TTGTTCAACC TGGCATAGCA TCTGAATTGG TCATCCCAAG CGAGCGCCTT	420
CACTACCGCA ATCAAGGTTG GCGCTCGGTT GAGACATCTG GTGTTGCTGA GGAGGAAGCC	480
ACCTCCGGTC TTGTCATGTT ATGCATACAT GGCTCTCCAG TTAACCTCCTA TACCAATACC	540
CCTTATACCG GTGCCCTTGG CTTACTGGAC TTTGCCTTAG AGCTTGAGTT TCGCAATCTC	600
ACCACCTGTA ACACCAATAC ACGTGTGTCC CGTTACTCCA GCACGGCCCCG TCACCGGCTC	660

40

CGCCGAGGGG CCGACGGGAC TCGGAGCTG ACCACAACCTG CAGCCACCAG GTTCATGAAA 720

GATCTCCACT TTACCGGCCT TAATGGGGTA GGTGAAGTCG GCCGCGGGAT AGCTCTAACA 780

5 TTACTTAACC TTGCTGACAC GCTCCTCGGC GGGCTCCCGA CAGAATTAAT TTCGTCGGCT 840

GGCGGGCAAC TGTTTTATTC CCGCCCGGTT GTCTCAGCCA ATGGCGAGCC AACCGTGAAG 900

CTCTATACAT CAGTGGAGAA TGCTCAGCAG GATAAGGGTG TTGCTATCCC CCACGATATC 960

10 GATCTTGGTG ATTCGCGTGT GGTCATTGAG GATTATGACA ACCAGCATGA GCAGGATCGG 1020

CCCACCCCGT CGCCTGCGCC ATCTCGGCCT TTTTCTGTTC TCCGAGCAAA TGATGTACTT 1080

15 TGGCTGTCCC TCACTGCAGC CGAGTATGAC CAGTCCACTT ACGGGTCGTC AACTGGCCCCG 1140

GTTTATATCT CGGACAGCGT GACTTTGGTG AATGTTGCGA CTGGCGCGCA GGCCGTAGCC 1200

CGATCGCTTG ACTGGTCCAA AGTCACCCTC GACGGGCGGC CCCTCCCGAC TGTTGAGCAA 1260

20 TATTCCAAGA CATTCTTTGT GCTCCCCCTT CGTGGCAAGC TCTCCTTTTG GGAGGCCGGC 1320

ACAACAAAAG CAGGTTATCC TTATAATTAT AATACTACTG CTAGTGACCA GATTCTGATT 1380

25 GAAAATGCTG CCGGCCATCG GGTCGCCATT TCAACCTATA CCACCAGGCT TGGGGCCGGT 1440

CCGGTCGCCA TTTCTGCGGC CGCGGTTTTG GCTCCACGCT CCGCCCTGGC TCTGCTGGAG 1500

GATACTTTTG ATTATCCGGG GCGGGCGCAC ACATTTGATG ACTTCTGCCC TGAATGCCGC 1560

30 GCTTTAGGCC TCCAGGGTTG TGCTTTCCAG TCAACTGTCG CTGAGCTCCA GCGCCTTAAA 1620

GTTAAGGTGG GTAAAACTCG GGAGTTG 1647

35 (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 984 base pairs
- (B) TYPE: nucleic acid
- 40 (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

5 (vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3
region

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

	GGTGCGGACG GGA	CTGCCGA GCTCACCACC	ACGGCTGCTA CCCGCTTTAT	GAAGGACCTC	60
	TATTTTACTA GTACTAATGG	TGTCGGTGAG ATCGGCCGCG	GGATAGCCCT CACCCTGTTC		120
15	AACCTTGCTG A	CACTCTGCT TGGCGGCCTG	CCGACAGAAT TGATTTCGTC	GGCTGGTGGC	180
	CAGCTGTTCT ACTCCCGTCC	CGTTGTCTCA GCCAATGGCG	AGCCGACTGT TAAGTTGTAT		240
20	ACATCTGTAG AGAATGCTCA	GCAGGATAAG GGTATTGCAA	TCCCGCATGA CATTGACCTC		300
	GGAGAATCTC GTGTGGTTAT	TCAGGATTAT GATAACCAAC	ATGAACAAGA TCGGCCGACG		360
	CCTTCTCCAG CCCCATCGCG	CCCTTTCTCT GTCCTTCGAG	CTAATGATGT GCTTTGGCTC		420
25	TCTCTCACCG CTGCCGAGTA	TGACCAGTCC ACTTATGGCT	CTTCGACTGG CCCAGTTTAT		480
	GTTTCTGACT CTGTGACCTT	GGTTAATGTT GCGACCGGCG	CGCAGGCCGT TGCCCGGTCTG		540
30	CTCGATTGGA CCAAGGTCAC	ACTTGACGGT CGCCCCCTCT	CCACCATCCA GCAGTACTCG		600
	AAGACCTTCT TTGTCCTGCC	GCTCCGCGGT AAGCTCTCTT	TCTGGGAGGC AGGCACAAC		660
	AAAGCCGGGT ACCCTTATAA	TTATAACACC ACTGCTAGCG	ACCAACTGCT TGTCGAGAAT		720
35	GCCGCCGGGC ACCGGGTCTG	C TATTTCCACT TACACCACTA	GCCTGGGTGC TGGTCCCGTC		780
	TCCATTTCTG CGGTTGCCGT	TTTAGCCCCC CACTCTGCGC	TAGCATTGCT TGAGGATACC		840
40	TTGGACTACC CTGCCCCGCG	C CCATACTTTT GATGATTTCT	GCCCAGAGTG CCGCCCCCTT		900

GGCCTTCAGG GCTGCGCTTT CCAGTCTACT GTCGCTGAGC TTCAGCGCCT TAAGATGAAG 960
 GTGGGTAAAA CTCGGGAGTT GTAG 984

5 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 984 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: Hepatits E Virus (Mexico strain) SG3
 region

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGGCCGACG GGA CTGCGGA GCTGACCACA ACTGCAGCCA CCAGGTTTCAT GAAAGATCTC 60
 25 CACTTTACCG GCCTTAATGG GGTAGGTGAA GTCGGCCGCG GGATAGCTCT AACATTACTT 120
 AACCTTGCTG ACACGCTCCT CGGCGGGCTC CCGACAGAAT TAATTTGTCG GGCTGGCGGG 180
 30 CAACTGTTTT ATTCCCGCCC GGTGTCTCA GCCAATGGCG AGCCAACCGT GAAGCTCTAT 240
 ACATCAGTGG AGAATGCTCA GCAGGATAAG GGTGTTGCTA TCCCCACGA TATCGATCTT 300
 GGTGATTCGC GTGTGGTCAT TCAGGATTAT GACAACCAGC ATGAGCAGGA TCGGCCACCC 360
 35 CCGTCGCCTG CGCCATCTCG GCCTTTTTTCT GTTCTCCGAG CAAATGATGT ACTTTGGCTG 420
 TCCCTCACTG CAGCCGAGTA TGACCAGTCC ACTTACGGGT CGTCAACTGG CCCGGTTTAT 480
 40 ATCTCGGACA GCGTGACTTT GGTGAATGTT GCGACTGGCG CGCAGGCCGT AGCCCGATCG 540

CTTGACTGGT CCAAAGTCAC CCTCGACGGG CGGCCCTCC CGACTGTTGA GCAATATTCC 600
 AAGACATTCT TTGTGCTCCC CCTTCGTGGC AAGCTCTCCT TTTGGGAGGC CGGCACAACA 660
 5 AAAGCAGGTT ATCCTTATAA TTATAATACT ACTGCTAGTG ACCAGATTCT GATTGAAAAT 720
 GCTGCCGGCC ATCGGGTCGC CATTTC AACC TATACCACCA GGCTTGGGGC CGGTCCGGTC 780
 10 GCCATTCTG CGGCCGCGGT TTTGGCTCCA CGCTCCGCC TGGCTCTGCT GGAGGATACT 840
 TTTGATTATC CGGGGCGGGC GCACACATTT GATGACTTCT GCCCTGAATG CCGCGCTTTA 900
 GGCCTCCAGG GTTGTGCTTT CCAGTCAACT GTCGCTGAGC TCCAGCGCCT TAAAGTTAAG 960
 15 GTGGGTAAAA CTCGGGAGTT GTAG 984

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
 406.3-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ACCTTGGA CT ACCCTGCCCC CGCCCATACT TTTGATGATT TCTGCCCAGA GTGCCGCCCC 60
 CTTGGCCTTC AGGGCTGCGC TTTCCAGTCT ACTGTCGCTG AGCTTCAGCG CCTTAAGATG 120
 40 AAGGTGGGTA AAAC TCGGGA GTTGTAG 147

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
406.3-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACTTTTGATT ATCCGGGGCG GCGCACACA TTTGATGACT TCTGCCCTGA ATGCCGCGCT 60
TTAGGCCTCC AGGGTTGTGC TTTCCAGTCA ACTGTCGCTG AGCTCCAGCG CCTTAAAGTT 120
AAGGTGGGTA AAACTCGGGA GTTGTAG 147

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma Strain)
ORF-3 region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

5 ATGAATAACA TGTCTTTTGC TGCGCCCATG GGTTCGCGAC CATGCGCCCT CGGCCTATTT 60
 TGTGCTGCT CCTCATGTTT TTGCCTATGC TGCCCGCGCC ACCGCCCGGT CAGCCGTCTG 120
 GCCGCCGTCG TGGGCGGCGC AGCGGCGGTT CCGGCGGTGG TTTCTGGGGT GACCGGGTTG 180
 10 ATTCTCAGCC CTTGCAATC CCCTATATTC ATCCAACCAA CCCCTTCGCC CCCGATGTCA 240
 CCGCTGCGGC CGGGGCTGGA CCTCGTGTTT GCCAACCCGC CCGACCACTC GGCTCCGCTT 300
 GCGGTGACCA GGCCCAGCGC CCCGCCGTTG CCTCACGTCG TAGACCTACC ACAGCTGGGG 360
 15 CCGCGCCGCT AA 372

(2) INFORMATION FOR SEQ ID NO:10:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 25 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (vi) ORIGINAL SOURCE:
 30 (C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)
 ORF-3 region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

35 ATGAATAACA TGTGGTTTGC TGCGCCCATG GGTTCGCCAC CATGCGCCCT AGGCCTCTTT 60
 TGCTGTTGTT CCTCTTGTTT CTGCCTATGT TGCCCGCGCC ACCGACCGGT CAGCCGTCTG 120
 40 GCCGCCGTCG TGGGCGGCGC AGCGGCGGTA CCGGCGGTGG TTTCTGGGGT GACCGGGTTG 180

ATTCTCAGCC CTTGCAATC CCCTATATTC ATCCAACCAA CCCCTTTGCC CCAGACGTTG 240
 CCGCTGCGTC CGGGTCTGGA CCTCGCCTTC GCCAACCAGC CCGGCCACTT GGCTCCACTT 300
 5 GGCGAGATCA GGCCCAGCGC CCCTCCGCTG CCTCCCGTCG CCGACCTGCC ACAGCCGGGG 360
 CTGCGGCGCT GA 372

(2) INFORMATION FOR SEQ ID NO:11:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

20

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
 406.4-2 region

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCCAACCCGC CCGACCACTC GGCTCCGCTT GCGGTGACCA GGCCCAGCGC CCCGCCGTTG 60
 30 CCTCACGTCG TAGACCTACC ACAGCTGGGG CCGCGCCGC 99

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 99 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
406.4-2 region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

10 GCCAACCAGC CCGGCCACTT GGCTCCACTT GGCGAGATCA GGCCCAGCGC CCCTCCGCTG 60
CCTCCCGTCG CCGACCTGCC ACAGCCGGGG CTGCGGCGC 99

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 660 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
ORF-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Arg Pro Arg Pro Ile Leu Leu Leu Leu Met Phe Leu Pro Met
1 5 10 15

Leu Pro Ala Pro Pro Pro Gly Gln Pro Ser Gly Arg Arg Arg Gly Arg
20 25 30

Arg Ser Gly Gly Ser Gly Gly Gly Phe Trp Gly Asp Arg Val Asp Ser
35 40 45

Gln Pro Phe Ala Ile Pro Tyr Ile His Pro Thr Asn Pro Phe Ala Pro

60

5

10

15

20

25

Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile Ser
195 200 205

35

40

Gly Trp Arg Ser Val Glu Thr Ser Gly Val Ala Glu Glu Glu Ala Thr
260 265 270

	Ser	Gly	Leu	Val	Met	Leu	Cys	Ile	His	Gly	Ser	Leu	Val	Asn	Ser	Tyr	
				275						280				285			
5	Thr	Asn	Thr	Pro	Tyr	Thr	Gly	Ala	Leu	Gly	Leu	Leu	Asp	Phe	Ala	Leu	
		290					295					300					
	Glu	Leu	Glu	Phe	Arg	Asn	Leu	Thr	Pro	Gly	Asn	Thr	Asn	Thr	Arg	Val	
	305					310					315					320	
10	Ser	Arg	Tyr	Ser	Ser	Thr	Ala	Arg	His	Arg	Leu	Arg	Arg	Gly	Ala	Asp	
					325					330					335		
	Gly	Thr	Ala	Glu	Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	Lys	Asp	
				340					345					350			
15	Leu	Tyr	Phe	Thr	Ser	Thr	Asn	Gly	Val	Gly	Glu	Ile	Gly	Arg	Gly	Ile	
		355						360					365				
	Ala	Leu	Thr	Leu	Phe	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	Leu	Pro	
20		370					375					380					
	Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	Arg	Pro	
	385					390					395					400	
25	Val	Val	Ser	Ala	Asn	Gly	Glu	Pro	Thr	Val	Lys	Leu	Tyr	Thr	Ser	Val	
					405					410					415		
	Glu	Asn	Ala	Gln	Gln	Asp	Lys	Gly	Ile	Ala	Ile	Pro	His	Asp	Ile	Asp	
				420					425					430			
30	Leu	Gly	Glu	Ser	Arg	Val	Val	Ile	Gln	Asp	Tyr	Asp	Asn	Gln	His	Glu	
		435						440					445				
	Gln	Asp	Arg	Pro	Thr	Pro	Ser	Pro	Ala	Pro	Ser	Arg	Pro	Phe	Ser	Val	
35		450					455					460					
	Leu	Arg	Ala	Asn	Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	Ala	Glu	Tyr	
	465					470					475					480	
40	Asp	Gln	Ser	Thr	Tyr	Gly	Ser	Ser	Thr	Gly	Pro	Val	Tyr	Val	Ser	Asp	
					485					490						495	

Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala Arg
 500 505 510

5 Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro Leu Ser Thr
 515 520 525

Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly Lys
 530 535 540

10 Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr Asn
 545 550 555 560

15 Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala Ala Gly
 565 570 575

His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly Pro
 580 585 590

20 Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala Leu Ala
 595 600 605

Leu Leu Glu Asp Thr Leu Asp Tyr Pro Ala Arg Ala His Thr Phe Asp
 610 615 620

25 Asp Phe Cys Pro Glu Cys Arg Pro Leu Gly Leu Gln Gly Cys Ala Phe
 625 630 635 640

Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys
 645 650 655

30 Thr Arg Glu Leu
 660

(2) INFORMATION FOR SEQ ID NO:14:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 660 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- 40 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico Strain)

5

ORF-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

10	Met Arg Pro Arg Pro Leu Leu Leu Leu Phe Leu Leu Phe Leu Pro Met	1	5	10	15
	Leu Pro Ala Pro Pro Thr Gly Gln Pro Ser Gly Arg Arg Arg Gly Arg	20	25	30	
15	Arg Ser Gly Gly Thr Gly Gly Gly Phe Trp Gly Asp Arg Val Asp Ser	35	40	45	
	Gln Pro Phe Ala Ile Pro Tyr Ile His Pro Thr Asn Pro Phe Ala Pro	50	55	60	
20	Asp Val Ala Ala Ala Ser Gly Ser Gly Pro Arg Leu Arg Gln Pro Ala	65	70	75	80
	Arg Pro Leu Gly Ser Thr Trp Arg Asp Gln Ala Gln Arg Pro Ser Ala	85	90	95	
	Ala Ser Arg Arg Arg Pro Ala Thr Ala Gly Ala Ala Ala Leu Thr Ala	100	105	110	
30	Val Ala Pro Ala His Asp Thr Ser Pro Val Pro Asp Val Asp Ser Arg	115	120	125	
	Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu Thr	130	135	140	
35	Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro Leu	145	150	155	160
	Asn Pro Pro Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met Ala	165	170	175	
40					

Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr Ile
 180 185 190

5 Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile Ser
 195 200 205

Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp Met
 210 215 220

10 Asn Ser Ile Thr Ser Thr Asp Val Arg Ile Leu Val Gln Pro Gly Ile
 225 230 235 240

Ala Ser Glu Leu Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn Gln
 245 250 255

15 Gly Trp Arg Ser Val Glu Thr Ser Gly Val Ala Glu Glu Glu Ala Thr
 260 265 270

20 Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Pro Val Asn Ser Tyr
 275 280 285

Thr Asn Thr Pro Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala Leu
 290 295 300

25 Glu Leu Glu Phe Arg Asn Leu Thr Thr Cys Asn Thr Asn Thr Arg Val
 305 310 315 320

Ser Arg Tyr Ser Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala Asp
 325 330 335

30 Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys Asp
 340 345 350

35 Leu His Phe Thr Gly Leu Asn Gly Val Gly Glu Val Gly Arg Gly Ile
 355 360 365

Ala Leu Thr Leu Leu Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu Pro
 370 375 380

40 Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg Pro
 385 390 395 400

00769066-044

Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser Val
 405 410 415

5 Glu Asn Ala Gln Gln Asp Lys Gly Val Ala Ile Pro His Asp Ile Asp
 420 425 430

Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His Glu
 435 440 445

10 Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser Val
 450 455 460

Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu Tyr
 465 470 475 480

15 Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Ile Ser Asp
 485 490 495

20 Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala Arg
 500 505 510

Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Pro Thr
 515 520 525

25 Val Glu Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly Lys
 530 535 540

Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr Asn
 545 550 555 560

30 Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala Gly
 565 570 575

35 His Arg Val Ala Ile Ser Thr Tyr Thr Thr Arg Leu Gly Ala Gly Pro
 580 585 590

Val Ala Ile Ser Ala Ala Ala Val Leu Ala Pro Arg Ser Ala Leu Ala
 595 600 605

40 Leu Leu Glu Asp Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe Asp
 610 615 620

Asp Phe Cys Pro Glu Cys Arg Ala Leu Gly Leu Gln Gly Cys Ala Phe
 625 630 635 640

Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys Val Lys Val Gly Lys
 5 645 650 655

Thr Arg Glu Leu
 660

10 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 549 amino acids

(B) TYPE: amino acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
 r62kDa, FIGURE 4

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Val Ala Pro Ala His Asp Thr Pro Pro Val Pro Asp Val Asp Ser
 30 1 5 10 15

Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu
 20 25 30

35 Thr Ser Ser Val Ala Thr Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro
 35 40 45

Leu Ser Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met
 50 55 60

40

Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr

	65		70		75		80									
	Ile	Arg	Tyr	Arg	Pro	Leu	Val	Pro	Asn	Ala	Val	Gly	Gly	Tyr	Ala	Ile
					85					90					95	
5																
	Ser	Ile	Ser	Phe	Trp	Pro	Gln	Thr	Thr	Thr	Thr	Pro	Thr	Ser	Val	Asp
					100				105					110		
10	Met	Asn	Ser	Ile	Thr	Ser	Thr	Asp	Val	Arg	Ile	Leu	Val	Gln	Pro	Gly
			115					120					125			
	Ile	Ala	Ser	Glu	Leu	Val	Ile	Pro	Ser	Glu	Arg	Leu	His	Tyr	Arg	Asn
		130					135					140				
15																
	Gln	Gly	Trp	Arg	Ser	Val	Glu	Thr	Ser	Gly	Val	Ala	Glu	Glu	Glu	Ala
	145					150					155					160
	Thr	Ser	Gly	Leu	Val	Met	Leu	Cys	Ile	His	Gly	Ser	Leu	Val	Asn	Ser
					165					170					175	
20																
	Tyr	Thr	Asn	Thr	Pro	Tyr	Thr	Gly	Ala	Leu	Gly	Leu	Leu	Asp	Phe	Ala
					180				185					190		
	Leu	Glu	Leu	Glu	Phe	Arg	Asn	Leu	Thr	Pro	Gly	Asn	Thr	Asn	Thr	Arg
25			195					200					205			
	Val	Ser	Arg	Tyr	Ser	Ser	Thr	Ala	Arg	His	Arg	Leu	Arg	Arg	Gly	Ala
		210					215					220				
30	Asp	Gly	Thr	Ala	Glu	Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	Lys
	225					230					235					240
	Asp	Leu	Tyr	Phe	Thr	Ser	Thr	Asn	Gly	Val	Gly	Glu	Ile	Gly	Arg	Gly
					245					250					255	
35																
	Ile	Ala	Leu	Thr	Leu	Phe	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	Leu
					260				265					270		
40	Pro	Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	Arg
					275			280					285			

	Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser
	290 295 300
5	Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile
	305 310 315 320
	Asp Leu Gly Glu Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His
	325 330 335
10	Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser
	340 345 350
	Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu
	355 360 365
15	Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Val Ser
	370 375 380
	Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala
20	385 390 395 400
	Arg Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro Leu Ser
	405 410 415
25	Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
	420 425 430
	Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
	435 440 445
30	Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala Ala
	450 455 460
	Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly
35	465 470 475 480
	Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala Leu
	485 490 495
40	Ala Leu Leu Glu Asp Thr Leu Asp Tyr Pro Ala Arg Ala His Thr Phe
	500 505 510

Asp Asp Phe Cys Pro Glu Cys Arg Pro Leu Gly Leu Gln Gly Cys Ala
515 520 525

5 Phe Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly
530 535 540

Lys Thr Arg Glu Leu
545

10 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 549 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- 25 (C) INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
r62kDa, FIGURE 4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

30 Ala Val Ala Pro Ala His Asp Thr Ser Pro Val Pro Asp Val Asp Ser
1 5 10 15

Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu
20 25 30

35 Thr Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro
35 40 45

Leu Asn Pro Pro Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met
50 55 60

40 Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr

	65					70						75					80
	Ile	Arg	Tyr	Arg	Pro	Leu	Val	Pro	Asn	Ala	Val	Gly	Gly	Tyr	Ala	Ile	
					85					90					95		
5																	
	Ser	Ile	Ser	Phe	Trp	Pro	Gln	Thr	Thr	Thr	Thr	Pro	Thr	Ser	Val	Asp	
				100					105					110			
10	Met	Asn	Ser	Ile	Thr	Ser	Thr	Asp	Val	Arg	Ile	Leu	Val	Gln	Pro	Gly	
			115					120					125				
	Ile	Ala	Ser	Glu	Leu	Val	Ile	Pro	Ser	Glu	Arg	Leu	His	Tyr	Arg	Asn	
		130					135					140					
15	Gln	Gly	Trp	Arg	Ser	Val	Glu	Thr	Ser	Gly	Val	Ala	Glu	Glu	Glu	Ala	
	145					150					155					160	
	Thr	Ser	Gly	Leu	Val	Met	Leu	Cys	Ile	His	Gly	Ser	Pro	Val	Asn	Ser	
				165						170					175		
20																	
	Tyr	Thr	Asn	Thr	Pro	Tyr	Thr	Gly	Ala	Leu	Gly	Leu	Leu	Asp	Phe	Ala	
			180						185					190			
	Leu	Glu	Leu	Glu	Phe	Arg	Asn	Leu	Thr	Thr	Cys	Asn	Thr	Asn	Thr	Arg	
25			195					200					205				
	Val	Ser	Arg	Tyr	Ser	Ser	Thr	Ala	Arg	His	Arg	Leu	Arg	Arg	Gly	Ala	
		210					215					220					
30	Asp	Gly	Thr	Ala	Glu	Leu	Thr	Thr	Thr	Ala	Ala	Thr	Arg	Phe	Met	Lys	
	225					230					235					240	
	Asp	Leu	His	Phe	Thr	Gly	Leu	Asn	Gly	Val	Gly	Glu	Val	Gly	Arg	Gly	
				245						250					255		
35																	
	Ile	Ala	Leu	Thr	Leu	Leu	Asn	Leu	Ala	Asp	Thr	Leu	Leu	Gly	Gly	Leu	
			260						265					270			
	Pro	Thr	Glu	Leu	Ile	Ser	Ser	Ala	Gly	Gly	Gln	Leu	Phe	Tyr	Ser	Arg	
40			275					280					285				

	Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser
	290 295 300
5	Val Glu Asn Ala Gln Gln Asp Lys Gly Val Ala Ile Pro His Asp Ile
	305 310 315 320
	Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His
	325 330 335
10	Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser
	340 345 350
	Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu
	355 360 365
15	Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Ile Ser
	370 375 380
	Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala
	385 390 395 400
	Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Pro
	405 410 415
25	Thr Val Glu Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
	420 425 430
	Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
	435 440 445
30	Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala
	450 455 460
	Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Arg Leu Gly Ala Gly
	465 470 475 480
	Pro Val Ala Ile Ser Ala Ala Ala Val Leu Ala Pro Arg Ser Ala Leu
	485 490 495
40	Ala Leu Leu Glu Asp Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe
	500 505 510

5 Phe Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys Val Lys Val Gly
530 535 540

10 (2) INFORMATION FOR SEQ ID NO:17:

15 (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3
region

Gly Ala Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe
1 5 10 15

Met Lys Asp Leu Tyr Phe Thr Ser Thr Asn Gly Val Gly Glu Ile Gly
20 25 30

Arg Gly Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly
35 40 45

Gly Leu Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr
50 55 60

40 Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr
65 70 75 80

	Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His	
	85	90 95
5	Asp Ile Asp Leu Gly Glu Ser Arg Val Val Ile Gln Asp Tyr Asp Asn	
	100	105 110
	Gln His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro	
	115	120 125
10	Phe Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala	
	130	135 140
	Ala Glu Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr	
	145	150 155 160
15	Val Ser Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala	
	165	170 175
	Val Ala Arg Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro	
20	180	185 190
	Leu Ser Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu	
	195	200 205
25	Arg Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr	
	210	215 220
	Pro Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn	
	225	230 235 240
30	Ala Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly	
	245	250 255
	Ala Gly Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser	
35	260	265 270
	Ala Leu Ala Leu Leu Glu Asp Thr Leu Asp Tyr Pro Ala Arg Ala His	
	275	280 285
40	Thr Phe Asp Asp Phe Cys Pro Glu Cys Arg Pro Leu Gly Leu Gln Gly	
	290	295 300

Val Gly Lys Thr Arg Glu Leu
5 325

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 327 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(vi) ORIGINAL SOURCE:
(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
SG3 region

25 Gly Ala Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe
1 5 10 15

30

35

Ser Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr
65 70 75 80

40 Thr Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Val Ala Ile Pro His
 85 90 95

	Asp	Ile	Asp	Leu	Gly	Asp	Ser	Arg	Val	Val	Ile	Gln	Asp	Tyr	Asp	Asn	
				100					105						110		
5	Gln	His	Glu	Gln	Asp	Arg	Pro	Thr	Pro	Ser	Pro	Ala	Pro	Ser	Arg	Pro	
			115					120					125				
	Phe	Ser	Val	Leu	Arg	Ala	Asn	Asp	Val	Leu	Trp	Leu	Ser	Leu	Thr	Ala	
		130					135					140					
10	Ala	Glu	Tyr	Asp	Gln	Ser	Thr	Tyr	Gly	Ser	Ser	Thr	Gly	Pro	Val	Tyr	
	145					150					155					160	
	Ile	Ser	Asp	Ser	Val	Thr	Leu	Val	Asn	Val	Ala	Thr	Gly	Ala	Gln	Ala	
					165					170					175		
15	Val	Ala	Arg	Ser	Leu	Asp	Trp	Ser	Lys	Val	Thr	Leu	Asp	Gly	Arg	Pro	
				180					185					190			
	Leu	Pro	Thr	Val	Glu	Gln	Tyr	Ser	Lys	Thr	Phe	Phe	Val	Leu	Pro	Leu	
20			195					200					205				
	Arg	Gly	Lys	Leu	Ser	Phe	Trp	Glu	Ala	Gly	Thr	Thr	Lys	Ala	Gly	Tyr	
		210					215					220					
25	Pro	Tyr	Asn	Tyr	Asn	Thr	Thr	Ala	Ser	Asp	Gln	Ile	Leu	Ile	Glu	Asn	
	225					230					235					240	
	Ala	Ala	Gly	His	Arg	Val	Ala	Ile	Ser	Thr	Tyr	Thr	Thr	Arg	Leu	Gly	
					245					250					255		
30	Ala	Gly	Pro	Val	Ala	Ile	Ser	Ala	Ala	Ala	Val	Leu	Ala	Pro	Arg	Ser	
				260					265					270			
	Ala	Leu	Ala	Leu	Leu	Glu	Asp	Thr	Phe	Asp	Tyr	Pro	Gly	Arg	Ala	His	
35			275					280					285				
	Thr	Phe	Asp	Asp	Phe	Cys	Pro	Glu	Cys	Arg	Ala	Leu	Gly	Leu	Gln	Gly	
		290					295					300					
40	Cys	Ala	Phe	Gln	Ser	Thr	Val	Ala	Glu	Leu	Gln	Arg	Leu	Lys	Val	Lys	
	305					310					315					320	

Val Gly Lys Thr Arg Glu Leu
325

(2) INFORMATION FOR SEQ ID NO:19:

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

15

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
406.4-2

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Asn Pro Pro Asp His Ser Ala Pro Leu Gly Val Thr Arg Pro Ser
1 5 10 15

25

Ala Pro Pro Leu Pro His Val Val Asp Leu Pro Gln Leu Gly Pro Arg
20 25 30

Arg

30

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

40

(iii) HYPOTHETICAL: NO

10-10-2000

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
406.4-2

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Asn Gln Pro Gly His Leu Ala Pro Leu Gly Glu Ile Arg Pro Ser
1 5 10 15

10

Ala Pro Pro Leu Pro Pro Val Ala Asp Leu Pro Gln Pro Gly Leu Arg
 20 25 30

Arg

15

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

20

- (A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

25

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma Strain)
ORF-3

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Asn Asn Met Ser Phe Ala Ala Pro Met Gly Ser Arg Pro Cys Ala
1 5 10 15

35

Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys Pro
 20 25 30

40

Arg His Arg Pro Val Ser Arg Leu Ala Ala Val Val Gly Gly Ala Ala

5

20

25

(iii) HYPOTHETICAL: NO

30

35

40

Leu Gly Leu Phe Cys Cys Cys Ser Ser Cys Phe Cys Leu Cys Cys Pro
20 25 30

5

10

15

Leu Ala Pro Leu Gly Glu Ile Arg Pro Ser Ala Pro Pro Leu Pro Pro
100 105 110

(2) INFORMATION FOR SEQ ID NO:23:

(A) LENGTH: 48 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain)
406.3-2

40

Glu Cys Arg Pro Leu Gly Leu Gln Gly Cys Ala Phe Gln Ser Thr Val
20 25 30

Ala Glu Leu Gln Arg Leu Lys Met Lys Val Gly Lys Thr Arg Glu Leu
 35 40 45

5 (2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

15

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E Virus (Mexico strain)
 406.3-2 region

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe Asp Asp Phe Cys Pro
 1 5 10 15

25

Glu Cys Arg Ala Leu Gly Leu Gln Gly Cys Ala Phe Gln Ser Thr Val
 20 25 30

30

Ala Glu Leu Gln Arg Leu Lys Val Lys Val Gly Lys Thr Arg Glu Leu
 35 40 45

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 540 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: protein

FBI/DOJ LABORATORY

```

(vi) ORIGINAL SOURCE:
      (C) INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
      r62kDa, 58.1 kDa

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala Val Ala Pro Ala His Asp Thr Pro Pro Val Pro Asp Val Asp Ser
1          5          10          15

Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu
          20          25          30

Thr Ser Ser Val Ala Thr Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro
          35          40          45

Leu Ser Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met
20          50          55          60

Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr
65          70          75          80

Ile Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile
25          85          90          95

Ser Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp
          100          105          110

Met Asn Ser Ile Thr Ser Thr Asp Val Arg Ile Leu Val Gln Pro Gly
30          115          120          125

Ile Ala Ser Glu Leu Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn
35          130          135          140

Gln Gly Trp Arg Ser Val Glu Thr Ser Gly Val Ala Glu Glu Glu Ala
145          150          155          160

Thr Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Leu Val Asn Ser
40          165          170          175

```

	Tyr Thr Asn Thr Pro Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala
	180 185 190
5	Leu Glu Leu Glu Phe Arg Asn Leu Thr Pro Gly Asn Thr Asn Thr Arg
	195 200 205
	Val Ser Arg Tyr Ser Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala
	210 215 220
10	Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys
	225 230 235 240
	Asp Leu Tyr Phe Thr Ser Thr Asn Gly Val Gly Glu Ile Gly Arg Gly
	245 250 255
15	Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu
	260 265 270
	Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg
20	275 280 285
	Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser
	290 295 300
25	Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile
	305 310 315 320
	Asp Leu Gly Glu Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His
	325 330 335
30	Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser
	340 345 350
	Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu
35	355 360 365
	Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Val Ser
	370 375 380
40	Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala
	385 390 395 400

141

Arg Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro Leu Ser
405 410 415

5 Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
420 425 430

Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
435 440 445

10 Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala Ala
450 455 460

Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly
465 470 475 480

15 Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala Leu
485 490 495

20 Ala Leu Leu Glu Asp Thr Leu Asp Tyr Pro Ala Arg Ala His Thr Phe
500 505 510

Asp Asp Phe Cys Pro Glu Cys Arg Pro Leu Gly Leu Gln Gly Cys Ala
515 520 525

25 Phe Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys
530 535 540

(2) INFORMATION FOR SEQ ID NO:26:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 540 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- 40 (vi) ORIGINAL SOURCE:
(C) INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)

r62kDa, 58.1 kDa

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

5 Ala Val Ala Pro Ala His Asp Thr Ser Pro Val Pro Asp Val Asp Ser
1 5 10 15

Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu
20 25 30

10 Thr Ser Ser Val Ala Ser Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro
35 40 45

Leu Asn Pro Pro Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met
15 50 55 60

Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr
65 70 75 80

20 Ile Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile
85 90 95

Ser Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp
100 105 110

25 Met Asn Ser Ile Thr Ser Thr Asp Val Arg Ile Leu Val Gln Pro Gly
115 120 125

Ile Ala Ser Glu Leu Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn
30 130 135 140

Gln Gly Trp Arg Ser Val Glu Thr Ser Gly Val Ala Glu Glu Glu Ala
145 150 155 160

35 Thr Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Pro Val Asn Ser
165 170 175

Tyr Thr Asn Thr Pro Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala
180 185 190

40 Leu Glu Leu Glu Phe Arg Asn Leu Thr Thr Cys Asn Thr Asn Thr Arg

	195	200	205
	Val Ser Arg Tyr Ser Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala		
	210	215	220
5	Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys		
	225	230	235 240
10	Asp Leu His Phe Thr Gly Leu Asn Gly Val Gly Glu Val Gly Arg Gly		
	245	250	255
	Ile Ala Leu Thr Leu Leu Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu		
	260	265	270
15	Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg		
	275	280	285
	Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser		
	290	295	300
20	Val Glu Asn Ala Gln Gln Asp Lys Gly Val Ala Ile Pro His Asp Ile		
	305	310	315 320
	Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His		
25	325	330	335
	Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser		
	340	345	350
30	Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu		
	355	360	365
	Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Ile Ser		
	370	375	380
35	Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala		
	385	390	395 400
	Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Pro		
40	405	410	415

Thr Val Glu Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
 420 425 430

5 Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
 435 440 445

Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala
 450 455 460

10 Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Arg Leu Gly Ala Gly
 465 470 475 480

Pro Val Ala Ile Ser Ala Ala Ala Val Leu Ala Pro Arg Ser Ala Leu
 485 490 495

15 Ala Leu Leu Glu Asp Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe
 500 505 510

20 Asp Asp Phe Cys Pro Glu Cys Arg Ala Leu Gly Leu Gln Gly Cys Ala
 515 520 525

Phe Gln Ser Thr Val Ala Glu Leu Gln Arg Leu Lys
 530 535 540

25 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 525 amino acids

(B) TYPE: amino acid

30 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)
 r62kDa, 56.5 kDa

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

5	Ala Val Ala Pro Ala His Asp Thr Pro Pro Val Pro Asp Val Asp Ser	1	5	10	15
	Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu	20	25	30	
10	Thr Ser Ser Val Ala Thr Gly Thr Asn Leu Val Leu Tyr Ala Ala Pro	35	40	45	
	Leu Ser Pro Leu Leu Pro Leu Gln Asp Gly Thr Asn Thr His Ile Met	50	55	60	
15	Ala Thr Glu Ala Ser Asn Tyr Ala Gln Tyr Arg Val Ala Arg Ala Thr	65	70	75	80
	Ile Arg Tyr Arg Pro Leu Val Pro Asn Ala Val Gly Gly Tyr Ala Ile	85	90	95	
20	Ser Ile Ser Phe Trp Pro Gln Thr Thr Thr Thr Pro Thr Ser Val Asp	100	105	110	
	Met Asn Ser Ile Thr Ser Thr Asp Val Arg Ile Leu Val Gln Pro Gly	115	120	125	
25	Ile Ala Ser Glu Leu Val Ile Pro Ser Glu Arg Leu His Tyr Arg Asn	130	135	140	
30	Gln Gly Trp Arg Ser Val Glu Thr Ser Gly Val Ala Glu Glu Glu Ala	145	150	155	160
	Thr Ser Gly Leu Val Met Leu Cys Ile His Gly Ser Leu Val Asn Ser	165	170	175	
35	Tyr Thr Asn Thr Pro Tyr Thr Gly Ala Leu Gly Leu Leu Asp Phe Ala	180	185	190	
40	Leu Glu Leu Glu Phe Arg Asn Leu Thr Pro Gly Asn Thr Asn Thr Arg	195	200	205	

	Val Ser Arg Tyr Ser Ser Thr Ala Arg His Arg Leu Arg Arg Gly Ala
	210 215 220
5	Asp Gly Thr Ala Glu Leu Thr Thr Thr Ala Ala Thr Arg Phe Met Lys
	225 230 235 240
	Asp Leu Tyr Phe Thr Ser Thr Asn Gly Val Gly Glu Ile Gly Arg Gly
	245 250 255
10	Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu
	260 265 270
	Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg
	275 280 285
15	Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser
	290 295 300
20	Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp Ile
	305 310 315 320
	Asp Leu Gly Glu Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His
	325 330 335
25	Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser
	340 345 350
	Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu
	355 360 365
30	Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Val Ser
	370 375 380
	Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala
	385 390 395 400
35	Arg Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro Leu Ser
	405 410 415
40	Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
	420 425 430

Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
 435 440 445

5 Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala Ala
 450 455 460

Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala Gly
 465 470 475 480

10 Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala Leu
 485 490 495

15 Ala Leu Leu Glu Asp Thr Leu Asp Tyr Pro Ala Arg Ala His Thr Phe
 500 505 510

Asp Asp Phe Cys Pro Glu Cys Arg Pro Leu Gly Leu Gln
 515 520 525

(2) INFORMATION FOR SEQ ID NO:28:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 525 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

30 (vi) ORIGINAL SOURCE:
 (C) INDIVIDUAL ISOLATE: Hepatitis E virus (Mexico strain)
 r62kDa, 56.5 kDa

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
 Ala Val Ala Pro Ala His Asp Thr Ser Pro Val Pro Asp Val Asp Ser
 1 5 10 15

40 Arg Gly Ala Ile Leu Arg Arg Gln Tyr Asn Leu Ser Thr Ser Pro Leu
 20 25 30

bioRxiv preprint doi: <https://doi.org/10.1101/009260>; this version posted April 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5

10

15

20

25

30

35

40

Asp Leu His Phe Thr Gly Leu Asn Gly Val Gly Glu Val Gly Arg Gly
245 250 255

Ile Ala Leu Thr Leu Leu Asn Leu Ala Asp Thr Leu Leu Gly Gly Leu
 260 265 270

5 Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser Arg
 275 280 285

Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr Ser
 290 295 300

10 Val Glu Asn Ala Gln Gln Asp Lys Gly Val Ala Ile Pro His Asp Ile
 305 310 315 320

Asp Leu Gly Asp Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln His
 325 330 335

15 Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe Ser
 340 345 350

20 Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala Glu
 355 360 365

Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Ile Ser
 370 375 380

25 Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val Ala
 385 390 395 400

Arg Ser Leu Asp Trp Ser Lys Val Thr Leu Asp Gly Arg Pro Leu Pro
 405 410 415

30 Thr Val Glu Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg Gly
 420 425 430

35 Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro Tyr
 435 440 445

Asn Tyr Asn Thr Thr Ala Ser Asp Gln Ile Leu Ile Glu Asn Ala Ala
 450 455 460

40 Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Arg Leu Gly Ala Gly
 465 470 475 480

150

Pro Val Ala Ile Ser Ala Ala Ala Val Leu Ala Pro Arg Ser Ala Leu
485 490 495

5 Ala Leu Leu Glu Asp Thr Phe Asp Tyr Pro Gly Arg Ala His Thr Phe
500 505 510

Asp Asp Phe Cys Pro Glu Cys Arg Ala Leu Gly Leu Gln
515 520 525

10 (2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

20 (iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: HEV 5' primer

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGGGGGGATC CATATGGCGG TCGCTCCGGC CCATGACACC CCG 43

30

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

40

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(C) INDIVIDUAL ISOLATE: HEV 3' primer

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATTAGAAGCT TCCGTGGCCA TTATATG

27

10 (2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

20

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

25 (C) INDIVIDUAL ISOLATE: internally consistent sequence of
two HEV 406.4-2 antigens

(ix) FEATURE:

(A) NAME/KEY: Modified-site

30

(B) LOCATION: 3

(D) OTHER INFORMATION: /note= "where Xaa is Q or P"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

35

(B) LOCATION: 5

(D) OTHER INFORMATION: /note= "where Xaa is G or D"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

40

(B) LOCATION: 7

(D) OTHER INFORMATION: /note= "where Xaa is L or S"

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 12

(D) OTHER INFORMATION: /note= "where Xaa is E or V"

5

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 13

(D) OTHER INFORMATION: /note= "where Xaa is I or T"

10

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 22

(D) OTHER INFORMATION: /note= "where Xaa is P or H"

15

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 24

(D) OTHER INFORMATION: /note= "where Xaa is A or V"

20

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 29

(D) OTHER INFORMATION: /note= "where Xaa is P or L"

25

(ix) FEATURE:

(A) NAME/KEY: Modified-site

(B) LOCATION: 31

(D) OTHER INFORMATION: /note= "where Xaa is L or P"

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Ala Asn Xaa Pro Xaa His Xaa Ala Pro Leu Gly Xaa Xaa Arg Pro Ser
 1 5 10 15

Ala Pro Pro Leu Pro Xaa Val Xaa Asp Leu Pro Gln Xaa Gly Xaa Arg
 20 25 30

40 Arg